THE IMPACT OF AMBIGUOUS CATEGORICAL FEEDBACK ON LEARNING AND PERFORMANCE IN A MULTIPLE-CUE DECISION PROBLEM

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

By

Edward Anthony Bryla

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Operations Research

Georgia Institute of Technology

August, 1974

THE IMPACT OF AMBIGUOUS CATEGORICAL FEEDBACK ON LEARNING AND PERFORMANCE IN A MULTIPLE-CUE DECISION PROBLEM

Approved:

Terry Connolly, Chairman

Russell G. Heikes

Thomas L. Sadosky

Date Approved by Chairman 9/16/74

ACKNOWLEDGMENTS

The writer wishes to express his deep appreciation and gratitude to Dr. Terry Connolly for his patience, wisdom, and aid during the completion of this study. He also thanks those excellent people, Drs. Russ Heikes and Tom Sadosky, who served on the reading committee, for their efforts and comments. Lastly appreciation is expressed to Al Dorris, without whose programming aid and encouragement this study would have had considerable distress.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

There appears to be a large number of important decisions in organizational contexts in which an individual expert is required to use several sources of imperfect information in reaching his decision. Medical diagnosis is, perhaps, the classic example, though similar processes may be seen in areas such as personnel selection, performance evaluation, and project selection/resource allocation in Research and Development. The Brunswik lens model has proved to be a fruitful paradigm into such judgment processes.

This research is formulated within this paradigm, and examines the impact of ambiguous, categorical feedback on learning and performance in judgment tasks. Such feedback provides the subject with a verbal statement of how well he is performing the task (e.g. good/bad) rather than providing exact numerical feedback. An experiment is conducted which examines the effects of varying (1) the number of classes of feedback provided, and (2) the severity of the evaluative cut-points between classes.

The results of the study indicate that feedback of performance evaluations in a judgment task has no effect on the learning or performance of a subject. The findings fail to confirm the findings of an identical study on this subject and seem to contradict past conclusions in regard to the

performance of subjects in multiple-cue probability learning tasks.

These findings cast some light on central questions in performance evaluations associated to Management By Objectives and parallel issues associated with alternative academic grading schemes.

CHAPTER I

INTRODUCTION

Purpose

The purpose of this chapter is to define the problem which is investigated by this research and to identify historical and current environments for which the research is especially relevant.

The Judgment Problem

Judgment can be defined, very generally, as the act of coming to a conclusion. This research investigates the act of coming to a conclusion, by an individual, about an underlying variable (distal) from a set of information inputs (cues). It attempts to determine the effects of feedback of performance evaluations on an individual in a judgment task.

Knowledge of what it is a man does, what are the characteristics of his performance, and how he acquires these characteristics is vital. Most existing theories of the development of human resources are built upon the question of how people learn and grow (Schein, 1961).

This research manipulates the laboratory environment in such a way as to make a desired conclusion to a judgment task the only one which is acceptable to a subject. Therefore it is also vital to know what kind of influences are exerted in such an environment and to assess the results of such influences. By allowing the subject to choose the direction of change in his behavior the research closely parallels most existing organizations. Hence solution to real managerial problems may be determined (Schein, 1961).

Historical and Current Problem Areas

The problem defined in the first part of this chapter has many outstanding examples in the fields of medical diagnostics, personnel selection, and other decision making environments. Conceptual models which would seem to be capable of providing useful insights and theoretical justification in these processes have been virtually non-existent. Organizational behavioralists have tended to ignore analytical models, while operation researchers have either ignored or made restrictive assumptions about behavior (Freeland, 1973).

This lack of catholicism is well illustrated in the field of Research and Development Project Selection and Resource Allocation.

The characteristic problems of this field are: (1) difficulty in estimating the value of scientific and technological information for decision making purposes (Dean and Roepche, 1969) and (2) difficulty in estimating the value of project selection decision variables, especially for new project proposals. Managers are required to rely heavily on their judgment of the caliber of individual and organizational sources of information which are provided as inputs to the decision analysis (Brandenberg, 1966).

Although some researchers have suggested that behavioral data, regarding how people tend to err during estimation, be built into their models (Moore and Baker, 1969) the value of the developed models has been limited by existing information structures (Cetron, Martino, and Roepche, 1967). These existing methods have a common link in that they rely on sound subjective judgments and differ, basically, in the complexity of the data that a decision maker or judge assimilates, organizes, and evaluates (Baker et al., 1970).

More recent models have adopted interactive approaches but continue to lack the authenticity which would be available by a better understanding of the R&D environment and of the behavorial processes by which decision and information systems become implemented and adopted (Baker and Freeland, 1972).

A recent model by Freeland (1973) stresses that organizational decisions are a function of the organization's information structure which changes as feedback about alternatives becomes available. But his attempt to add behavorial realism to his model is restricted by the assumption that decision makers are unable to consider all the complexities which enter into the decision process, and rather rely on aggregate measures of performance (Freeland, 1973).

The utilization of feedback in a decision process is a basic component in the concept of Management By Objectives. This managerial development attempts to diminish the complex problems of communication of job related, risk taking information and treats the communication of goals and results as the primary communication problem. It attempts to provide information to the individual in the form of reports by which he can measure his own performance (Odiorne, 1965).

However, the evolution of MBO has failed to accomplish this. It has dysfunctionally proliferated forms; goal forms, skill inventory forms, progress review forms, etc. (Howell, 1970). Furthermore, the "make or break" element of the concept, the performance review has also proven to be dysfunctional e.g. The American Chain and Cable Company uses annual review sessions but in these sessions a subordinate is not told any actual ratings, rather he is given a "sense of the ratings". A manager never gets to see his appraisal (Hetland, 1973). At the operational level traditional MBO feedback has been meaningless for it is too cumbersome, vague, and unproductive (Duncan, 1973).

Although most of the literature in this field suggests that feedback has a strong effect on the success of MBO implementation and some evidence that frequency of feedback helps explain attitudes, performance, and aspiration levels (Cook, 1968) there are few other cases of empirical research in the field (Parrish, 1973).

Since the necessary knowledge to overcome these problems is lacking the success of MBO will continue to be impaired. By examining the effects of feedback of performance evaluation on the performance of an individual perhaps some answers will begin to be provided.

As illustrated more knowledge about the field of human cognitive activity is a necessity. In order to make a contribution to the broad realm of this activity this research will attempt to determine the effects of ambiguous, categorical feedback on the learning and performance of an individual in a judgment task.

The next chapter of this thesis will delineate the framework within which this research evolved, relate the results of past research in this field and outline the objectives of this thesis.

CHAPTER II

REVIEW OF THE LITERATURE

Purpose

The purpose of this chapter is to introduce the reader to the rather large body of literature applicable to this research, and to identify within this body of literature the specific approach that will be utilized in this research. It outlines the methodology within which the research will be conducted and defines exactly what aspects of human cognitive activity are to be studied.

Information Processing

The processing of information by an individual occurs at several levels. The concern of this research is solely with cognitive operations performed on symbols, signs, and facts.

Prior to the 1960's there was relatively little research accomplished on information processing at this level. Since 1960 there have been several hundred studies performed within the topics of information utilization in judgment and decision making (Slovic and Lichtenstein, 1971).

Research Approach

This past research has been accomplished within two basic schools, the "regression" and the "Bayesian" approaches.

Although these approaches have paralled each other to a great extent the latter school seems rather uninterested in learning with a few notable exceptions. Martin and Gettys (1968) compared performance with nominal feedback or probabilistic feedback; Phillips and Edwards (1966) studied performance under different payoff conditions; and the studies of Peterson, Ducharme, and Edwards (1968) and Wheeler and Beach (1968) "oriented to misperception explanation of conservatism" (Slovic and Lichtenstein, 1971). For this reason this research restricts itself to the regression approach.

Within this approach two paradigms have been utilized; the correlational paradigm and the ANOVA paradigm. The ANOVA paradigm restricts any research to that concerned with categorical, uncorrelated cues (Goldberg, 1968) hence this study will be within the framework of the correlational paradigm.

The research trends within the correlational paradigm have either focused on the individual or on the probabilistic intercorrelation between the individual and the environmental components of the judgmental situation. In order to best portray the real environment and permit the analysis of the relative contribution of environmental factors to an individual's behavior this research utilizes a conceptual model from which the latter focus developed.

The Lens Model

Egon Brunswik (1965) developed the lens model to represent the judgmental situation. He advocated the study of a judge in a realistic situation, in experiments representative of a real ecology. The lens model provides a means for appropriately specifying the structure of the situational variables in an experiment (Slovic and Lichtenstein, 1971).

The basic elements of this model are shown in Figure 1, and the following set of definitions and conceptual explanation.

r_a - achievement coefficient of the individual; the correlation between his prediction and the actual distal variable.

r_{ei} - the simple correlation between a single cue, x_i, and the distal variable.

r_{si} - the simple correlation between a single cue, x_i, and the individual's estimate.

$r_{x_i x_j}$ - the intercorrelation between the cue x_i and the cue x_j.

X_i - the ith cue or information source about the distal variable Y_e

Y_e - the true state of the distal variable

Y_s - the subject's estimate of the distal variable

Concept

An individual is conceived of as using the set of cues, x_i's, to make a prediction, Y_s, as to the true state of the distal variable, Y_e. Each cue is related to the distal variable in some way (the extent of this relationship is

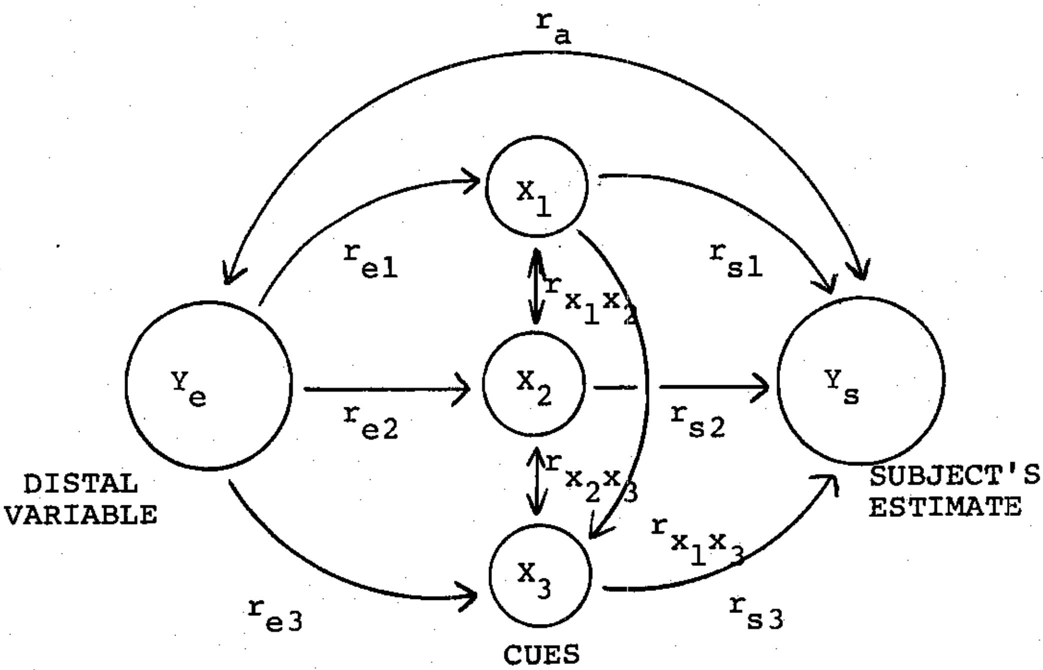

Figure 1. <u>Lens Model</u>

measured and denoted by the correlational coefficient r_{ei}) and is utilized by the individual to make his prediction in some way (the extent of this utilization is measured and denoted by the correlational coefficient r_{si}). The cues are also related to each other in some way (the extent of this relationship is measured and denoted by the cue intercorrelation $r_{x_i x_j}$). The individual's performance is measured by the performance index r_a (the correlation between Y_e and Y_s).

Brunswik's general argument has come to be accepted as scientifically respectable and a useful point of departure for valuable theoretical and empirical work. It argues that the environment is an uncertain, probabilistic one; that judgments are made on the basis of probabilistic data; that correlations are determined by the natural environment; and without a knowledge of the limitations placed on the subject by the statistical characteristics of the environment it is impossible to examine his behavior (Hursh, Hammond, and Hursh, 1964). Hammond (1960) noted that the lens model permits an analysis of the relative contribution of environmental factors to an individual's behavior (Slovic and Lichtenstein, 1971).

The basic problem in any such situation, where decisions are based upon evaluating and processing a set of cues or stimuli, is that the task is probabilistic in nature. This is because of the uncertainties in the relationships between cues, upon which judgment is based, and the distal or criterion variable, to be predicted (Castellan, 1973).

The conceptual framework of the Brunswik lens model was given a firm mathematical basis by Hursh, Hammond, and Hursh, in 1964. Tucker (1964) developed this work into a clearer and more useful model. His formulation is presented in Appendix A. These formulations have been utilized by the overwhelming majority of the past research work in the field. Castellan (1972) extended the basic model from that of a single distal variable, single response system to that of a multiple distal variable, multiple response system.

Hammond and Summers (1972) demonstrated that these formulations to the multiple-cue probability learning task might be insufficient. They hypothesized that the acquisition, and application of knowledge are independent components of learning in cognitive tasks. They showed that knowledge and control can be disentangled empirically and statistically. Further that even when knowledge is complete imperfect cognitive control can prevent high achievement.

Past Research Efforts

The use of multiple regression as a model for human use of information has been applied to such diverse topics as depth perception (Brunswik, 1956); person perception (Brunswik, 1956); clinical inference (Hammond, Hursh, and Todd, 1964); and conflict resolution (Hammond, 1965) (Beach, 1967). It assumes that men attempt to function effectively in an environment in which they receive only equivocal information (Beach, 1967). A major purpose within this field has been to determine how to optimize a subject's judgmental system (Beach, 1967). The foci of the research effort have been: (1) specification of the policy equation for the judge, including the question of whether or not non-linearities should be included; and (2) the learning or adaptation of the judge in an uncertain environment. Experimental work in the latter area has documented man's difficulty in processing multi-dimensional and probabilistic information. As Chapter I pointed out, these difficulties persist outside the artificial confines of the lab, even where familiar sources of information are used to make decisions (Slovic and Lichtenstein, 1971).

Multiple-Cue Probability Learning

The key to an expert's success resides in his ability to interpret and integrate information appropriately. This means that an individual must weigh items of information differentially according to their relevance and be able to qualify his interpretations of a given fact when other considerations make such qualifications necessary. The learning by an individual to utilize information has been categorized as a focal topic of the correlational paradigm. This learning is best defined as the process of learning to utilize complex configurations of ambiguous or probabilistic cues (Dudycha and Naylor, 1966). Multiple cue probability learning research has attempted to understand cognitive processes to learn why some judges are more accurate than others and to utilize this knowledge to train other judges (Slovic and Lichtenstein, 1971). This part of the chapter will specifically review the literature of this field.

Basic Findings

Smedslund (1955) demonstrated that subjects can learn to use cues appropriately and develop strategies in a probability learning stuation. Summers (1962) found that subjects could differentially respond to unequal cue validities (Dudycha and Naylor, 1966).

Uhl (1963) found that there are differential learning rates in tasks involving different cue validities. Brunswik (1956), Rommetveit (1960), and Rappaport (1963) demonstrated that there is a difference in multiple-cue inference tasks for subjects functioning under an analytic set as opposed to subjects functioning under an intuitive set. Wallach and Koger (1959) found a difference in performance in multiple-cue probability learning tasks between men and women (Todd and Hammond, 1965).

Peterson, Hammond, and Summers (1965) determined that subjects in these tasks fail to weight heavily enough the most valid of given cues, and slightly overweight those with the lowest validity. They also found, as did Summers (1969), that subjects can learn to detect changes in relative cue weight over time, although they do so slowly (Slovic and

Lichtenstein, 1971).

Todd and Hammond (1965) demonstrated that multiple-cue probability learning is slow and hypothesized that it was due to the type of feedback employed, outcome feedback.

Hammond and Summers (1965), and Summers (1967) found that the learning of non-linear relationships can occur but that it is slower and less effective than the learning of linear relationshps. Summers and Hammond (1966), Hammond and Summers (1968), and Earle (1970) found that this was especially true if the subjects were not forewarned that the relationships were non-linear (Slovic and Lichtenstein, 1971).

Several investigators have argued that judgment policies are frequently correct but are executed in an inconsistent manner and that the removal of all feedback increases consistency in response (Hammond and Summers, 1972).

Cue Research

The consistency of the cues also has been an issue of investigation. Slovic (1966) defined cue consistency as the agreement of all cues in their implication for the attributes being judged. He found that cue consistency influences the manner in which a person utilizes information when making a judgment, specifically that inconsistent cues are disregarded or not used by a subject. However he also found that an overriding consideration is the intrinsic validity felt for the cue by a subject. He hypothesized that this aspect is the determining factor in deciding whether a cue will be discounted or used in the event of cue inconsistency. Dudycha and Naylor (1966) demonstrated, in a two due task, that the pairing of a cue of low or medium validity with one of high validity is detrimental to performance, while pairing of a cue of low validity with another of medium or low validity is facilitative.

Summers (1967) found that subjects seem to learn which cue to use more easily than which functional rule relates known valid cues to the distal variable and that the learning of both simultaneously is especially difficult (Slovic and Lichtenstein, 1971). A related finding (Schenck and Naylor, 1968) showed that when cues are intercorrelated a subject's responses will become systematically more of a linear function of the cues as the intercorrelations of the cues increase. Brehmer (1970) showed that subjects can learn to use valid cues even when they are not perceived with perfect reliability (Slovic and Lichtenstein, 1971).

McMiller (1971) found that a cue is not a simple entity. He determined that subjects seem to follow one of two strategies; they either follow a verbal strategy suggested by the cue's label or a math strategy based on the numerical values of the cue-distal variable correlation. He suggested that this was a result of the difference between symbolic and semantic information. It is rather uncertain to what degree symbolic information is divorced from all meanings and what effect this relationship has.

Several investigators have found additional detrimental effect on subject performance resulting from a failure to take appropriate account for the redundancies among cues (Slovic and Lichtenstein, 1971).

The experimental work described above is excellently reviewed in Slovic and Lichtenstein (1971). It documents man's difficulty in processing multi-dimensional and probabilistic information. However, they remark that few variables have been explained in much depth, even such fundamental ones as the number of cues, cue redundancy or the effects or stress. A more recent study continues this pattern. Dudycha, Dimoff, and Dudycha (1973) attempted to break away from previous investigations which they felt focused only upon a man's ability to use probabilistic information in a static environment.

They varied the cue validities, ecological exposure length, and the amount of information available regarding the shift in validities. They studied and found that subjects are able to detect the dynamics of the environment and adjust their strategies in an appropriate direction. Despite this recognition they found that performance and consistency are depressed if the movement was from a low validity ecology to a high validity ecology. However performance in a low validity ecology is enhanced if preceded by a high validity ecology.

Feedback Research

As Slovic and Lichtenstein (1971) pointed out, few variables have been examined in much depth in multiple cue probability learning studies. This is especially true in an examination of the area of feedback research.

Feedback can be divided into three types: (1) outcome feedback, (2) structure of ecology feedback, and (3) subject performance feedback (Castellan, 1974). Most of the past research has followed one of the first two categories. There have been few studies which have examined the effects of feedback which is delayed, intermittent, error-prone or reduced to an evaluative good/bad dimension (Connolly and Miklausich, 1974).

Classical probability studies have relied heavily upon outcome feedback as the information source to the subject about the task. While outcome feedback does provide the subject with an indication of how accurate his prediction is on each trial, it is implicitly assumed that the subject has a memory capacity sufficient to keep an appropriate record of the relevant aspects of his performance, which are necessary for him to predict well on subsequent trials (Castellan, 1974).

Todd and Hammond (1965) had hypothesized that multiple-cue probability learning was slow under outcome feedback and that it is the kind of feedback information utilized that is important and not the quantity.

Findings such as this had led researchers to develop the other type of traditional feedback, lens model. This formulation utilizes cue utilization information, cue validity information, or a combination of the two.

Todd and Hammond (1965), Newton (1965), and later Hammond and Boyle (1970) demonstrated that this type of feedback enhances performance in comparison to outcome feedback (Slovic and Lichtenstein, 1971). Summers and Hammond (1966) found the same result in a task which had cues which were non-linearly related to the distal variable. Magnusson and Nystedt (1966) determined that cue validity feedback is better when the information was presented in the form of graphic or verbal descriptions rather than merely as the product moment correlation coefficients.

The general thrust of the studies reported to this point, both outcome feedback type and lens model feedback type, has been that outcome feedback produces rather slow learning for all but the simplest of task structures and that at a certain cost in time and effort lens model feedback will create a situation more conducive to learning than outcome feedback (Connolly and Miklausich, 1974).

However in the real world one does not always discover the right answer after making a judgment, nor is it likely that the feedback will be in the form of cue validities or cue utilization coefficients. Even in the lab setting this requires extensive experimentor intervention in analyzing and presenting information to the subject. More likely it will be in some crude form (Connolly and Miklausich, 1974). As pointed out in the first chapter there seem to be a whole set of decision makers who utilize multiple information systems and receive neither ecological information nor the exact information regarding the task.

Bjorkman (1972) and Castellan (1974) have pointed out the importance and role of different types of information in judgment tasks. Yet relatively few studies have examined the effects of different sorts of feedback upon performance.

One such study however has recently been completed. Miklausich (1973) examined the effects of error-prone outcome feedback. He found that performance is degraded as the amount or error in the feedback increases. He also found that the degradation was primarily due to a failure of the subjects' estimates to fit the linear regression equation of their estimates, rather than due to the failure of the linear regression equation estimates of the subjects' predictions of the outcome variable to correlate with the linear regression equation estimates of the outcome variable values.

Additionally there are two reported studies which can be classified as utilizing subject performance feedback. Castellan (1974), as part of a more traditional study, included feedback which consisted of percentage of correct responses across trials. He varied this feedback by giving it for the last twenty (20) trials undertaken, all previous trials undertaken, or both. He found that the percentage of correct answer feedback enhanced performance only across tasks with a single relevant cue. He also found that both delayed and intermittent feedback are suppressors of performance regardless of the type of feedback utilized. However if both were employed performance was enhanced in comparison to traditional outcome feedback. It must be noted that Castellan's study was for a non-metric case in which the subjects were asked to estimate the correct distal variable from a set of two. He noted that in metric ecologies no consistent results have emerged.

The only other study of this type was conducted by Rose (1974). He placed his subjects in the position of a person trying to predict the future performance of the stock market. He provided the opinions of three imaginary experts, stockbrokers, as to what the daily stock market change would be. Feedback was supplied in the form of a categorical evaluation, eg. Good/Bad. The categories ranged from two classes to five. His principal findings were that not only did the subjects perform worse with the evaluative feedback than with no feedback at all but also that their performance was degraded as the number of classes of evaluations increased. The feedback seemed to hinder the subject's application of his own regression equation.

Although it has generally been concluded in past studies that performance is effected by the loss of cognitive control, the fact that the increased refinement of information was a contributing factor in the degradation of performance is stunning. Rose suggested that this result is possibly the result of the severity of his evaluative cutpoints which determined the feedback. It seems more likely that the results were due to the experimental structure. As the number of feedback categories increased the cutpoints were not controlled across categories. Thus the performance of a subject in the five class category could have been as good or better than that of someone in the two class category and yet he would receive misleading, unrepresentative, and incorrect evaluations as feedback.

Research Questions

As pointed out a key consideration in Brunswikian type studies is the close resemblance of an experimental design to the real ecology. Examples from what seems to be the real world of the decision maker such as: workers in the business world whose feedback is whether or not they get a pay raise; students who receive a pass/fail or some form of letter grade; Department of Health, Education and Welfare progress review reports (Brady, 1973), as well as those situations covered in Chapter I; seem to call for feedback of the type utilized by Rose in his study.

Additionally it has been hypothesized in the Management By Objectives that feedback influences attitudes and performance results (Cook, 1968) and that good results should be reinforced by feedback of success while failures should be used only as a platform for coaching (Odiorne, 1965).

Rose hypothesized that feedback in the form of evaluations is detrimental to the performance of a subject in multiple-cue probability tasks as the number of categories of evaluations is increased. As noted above this effect may be due to the severity of the cutpoints he used in determining the evaluations.

Ambiguous, categorical feedback is an extention of Rose's definition of evaluative feedback. It provides the subject information on how well he is performing rather than revealing some numerical feedback. However in order to fully explore the question different levels of the severity of the cutpoints are utilized and controlled across groups of subjects.

This research will attempt to extend Rose's study by addressing the following questions:

What effect does the use of categorical, ambiguous feedback have upon the performance of an individual in a multiple-cue probability learning task?

<u>Hypothesis</u>

Performance will be effected as the number of categories of ambiguous, categorical feedback increases.

Performance will be effected as the harshness of the evaluations of the subject's performance decreases.

Performance will be effected by the interaction of the severity of the evaluations and the number of categories.

Chapter III will deal with experimental design.

CHAPTER III

METHODS AND PROCEDURES

Purpose

The purpose of this chapter is to explain the experimental design considerations that were followed in this study and describe the methods that were used in the experiment.

Experimental Design Considerations

The design of this experiment had as its objective a reasonably accurate portrayal of a real ecology. Hence factors such as types of feedback, time pressure, verbal context, cue interrelationships, and possible earlier experiences of the subjects were considered together to form an internally consistent package. These factors, if properly dealt with, then form a task for the subjects in the experiment which is neither too easy (a task in which all subjects achieve and hold the same level of learning after only a few trials) nor too hard (a task in which no subject is able to display any learning despite the number of trials) (Rose, 1974). Additionally they serve to aid the subjects to "make sense of the task" (Connolly, 1972).

Verbal Context

The basic verbal context used in this experiment is almost an exact replication of that utilized by Rose (1974).

It placed the subjects in the position of estimating or predicting the future performance of the stock market. All subjects were given the opinions of three experts, stockbrokers, and were asked to make their estimates solely on this information. They were also told that they would receive, after their estimate, an indication of their performance either in the form of the actual market change or in the form of an evaluation.

Time Pressure

When subjects arrive at an estimate they can do so using either an analytical or an intuitive model. The analytical model is one where the subject is allowed to develop some formulation such that when given inputs for his formulation he can use external sources such as paper, pencil, calculator, or other computer. The intuitive mdoel does not allow the subject to use external sources. An intuitive model was imposed on all subjects in order to study their instinctive utilization of the feedback. The feedback was given to the subjects individually by circling the appropriate feedback entry on a response sheet, in the evaluative groups, to prevent subjects' performance from influencing other subjects. The outcome feedback groups received their feedback verbally. In each case the pace of the slowest subject determined when the feedback was given.

Task Properties

An experimental task representing the cue-distal variable relationship was constructed by means of computer programs (See Appendix B). The task consisted of 80 trials divided into two blocks of 40 (See Appendix C). The cue validities were:

Theoretical	Experimental
$r_{e1} = .554$	$r_{e1} = .677$
$r_{e2} = .894$	$r_{e2} = .900$
$r_{e3} = .707$	$r_{e3} = .704$

for the entire sequence of 80 trials. The cue intercorrelations were:

Trials 1-40	Trials 41-80
$r_{12} = .680$	$r_{12} = .572$
$r_{13} = .448$	$r_{13} = .345$
$r_{23} = .604$	$r_{23} = .696$

The multiple correlations between the cues and the distal variable were: .926 for trials 1-40 and .928 for trials 41-80. Thus the task provided a degree of probabilism or indeterminancy as well as cue consistency and congruency.

Method

Subjects: The subjects in the experiment were 50 graduate students at the Georgia Institute of Technology and also were all officers in the United States Army.

Apparatus: The subject task was to receive three (3) cues (the expected stock market performance from two relatively poor experts and one relatively good expert ie. stockbrokers) and on the basis of the cues to estimate the value of the distal variable, the actual performance of the stock market as indicated by the Dow Jones Industrial Average. The distal variable, Y_e, is a random sample of a Normal Distribution ($\mu = 0$, $\sigma^2 = 100$) such that 67% of its values were within ± 10 points of 0. The cues are:

$X_i = Y_e \pm e_i$, $i = 1,2,3$, where e_i is a random sample from a Normal Distribution, with μ_i's=0, and σ_1^2=225, σ_2^2=25, σ_3^2=100, ie., $e_1 \sim N(0,225)$; $e_2 \sim N(0,25)$; $e_3 \sim N(0,100)$.

Procedure

The cues were displayed to the subjects by means of a view graph (See Appendix C.) Additionally each trial was read out loud to them and they were asked to record the cues on response sheets (See Appendix D). An identical verbal set of instructions was read to all subjects (See Appendix E). The feedback was presented in one of two ways; for Groups I and II (outcome feedback) it was provided verbally, otherwise it was in the form of evaluations provided on their response sheets (See Table 1).

Table 1. Feedback for Evaluative Groups

Groups/Cutoffs (# of possible evaluations-severity level)	(absolute deviation from the correct answer; ie. $\lvert Y_e - Y_s \rvert$)					
	0	0<but<1	1≤but<2	2≤but<4	4≤but<8	≥8
III (2-harsh)	Good	Good	Good	Bad	Bad	Bad
IV (2-mild)	Good	Good	Good	Good	Bad	Bad
V (3-harsh)	V. Good	V. Good	Good	Bad	Bad	Bad
VI (3-mild)	V. Good	V. Good	V. Good	Good	Bad	Bad
VII (3-harsh)	Good	Good	Good	Bad	V. Bad	V. Bad
VIII (3-mild)	Good	Good	Good	Good	Bad	V. Bad
IX (4-harsh)	V. Good	V. Good	Good	Bad	V. Bad	V. Bad
X (4-mild)	V. Good	V. Good	V. Good	Good	Bad	V. Bad

Feedback

Within the task the subjects were divided equally into 10 feedback groups. All groups received an initial block of 40 trials without feedback. The second block of 40 provided different feedback to different groups. Groups I and II received outcome feedback (the correct answer). The feedback for Groups III-X is displayed, along with the evaluative cutpoints in Table 1.

Subjective Evaluation

At the conclusion of all 80 trials a questionnaire was given to all 50 subjects in order to gain some insight into their perception regarding the experiment and their performance. (See Appendix F).

The next chapter presents the results of the experiment.

CHAPTER IV

RESULTS

Purpose

The purpose of this chapter is to present the results of the experiment. The first section deals with the testing and analysis of the hypotheses of this research while the second and third sections deal with further analysis of the data.

Primary Analysis

The first hypothesis to be tested by this experiment was that: The performance of a subject in a multiple-cue probability learning task will be effected as the number of categories of ambiguous, categorical feedback given to him increase. The data collected in the experiment was analyzed using a computer program (See Appendix G). This program derived regression equations to fit the model of each subject's estimates and thereby provided the parameters $\hat{Y}_e$ and $\hat{Y}_s$, which in turn generated the correlations of concern namely, r_a, R_e, R_s, and G. Fisher's Z transformations were then applied to the correlations since the distribution of Z values is approximately normal (McNemar, 1949). This allowed averaging and subsequent analysis of variance.

An analysis of variance was performed to compare the achievement level reached by all groups. The data failed to confirm the hypothesis. It indicated that the number of categories of feedback had no significant effect on the performance of the subjects. Results and analysis are displayed in Figure 2 and Table 2. The values of r_a; R_e, limit of achievement; R_s, cognitive control; G, knowldege of the task; and r_{avg}, limit of achievement by a strict averaging of the cues were plotted to aid in the interpretation of the data.

The second hypothesis to be tested was that: The performance of a subject in a multiple-cue probability learning task will be effected as the harshness of the evaluative cutoff points decrease. Utilizing the same analytical techniques, again the hypothesis was not confirmed. The harshness of the cutoffs had no significant effect on performance. These results and analysis are also displayed in Figure 2 and Table 2.

The third hypothesis was also not confirmed in that the data indicated that the interaction between the number of categories and the cutoff point harshness had no significant effect on performance. These results and analysis are displayed in Figure 2 and Table 2.

Table 2. Mean r_a by Groups, Trials 41-80

Group/Mean r_a (based on Z transformations)	
Group I (outcome feedback)	.861427
Group II (outcome feedback)	.859157
Group III (2 cat.-harsh)	.853153

Group IV (2 cat. - mild)	.850058
Group V (3 cat. A - harsh)	.872221
Group VI (3 cat. A - mild)	.85625
Group VII (3 cat. B - harsh)	.8469
Group VIII (3 cat. B - mild)	.865646
Group IX (4 cat. - harsh)	.83315
Group X (4 cat. - mild)	.844644

all n = 5

ANOVA (trials 41-80) (based on Z transformations)

Source of Variation	d.f.	M.S.	F ratio
Number of Categories	3	.01511468	.7115073003
Harshness	1	.00076266	.035901399
N X H	3	.02567028	.4028003656
Error	32	.679781864	
Total	39	.75841879	

Secondary Analysis

In addition to the testing of the major hypotheses the data collection made possible the investigation of several other items of interest.

Two groups of the subjects in the experiment received traditional outcome feedback rather than any of the ambiguous, categorical types. A comparison of the mean of these groups performance was made against the mean of the other groups' performance. The hypothesis tested was that: There is no significant difference between the performance of the subjects

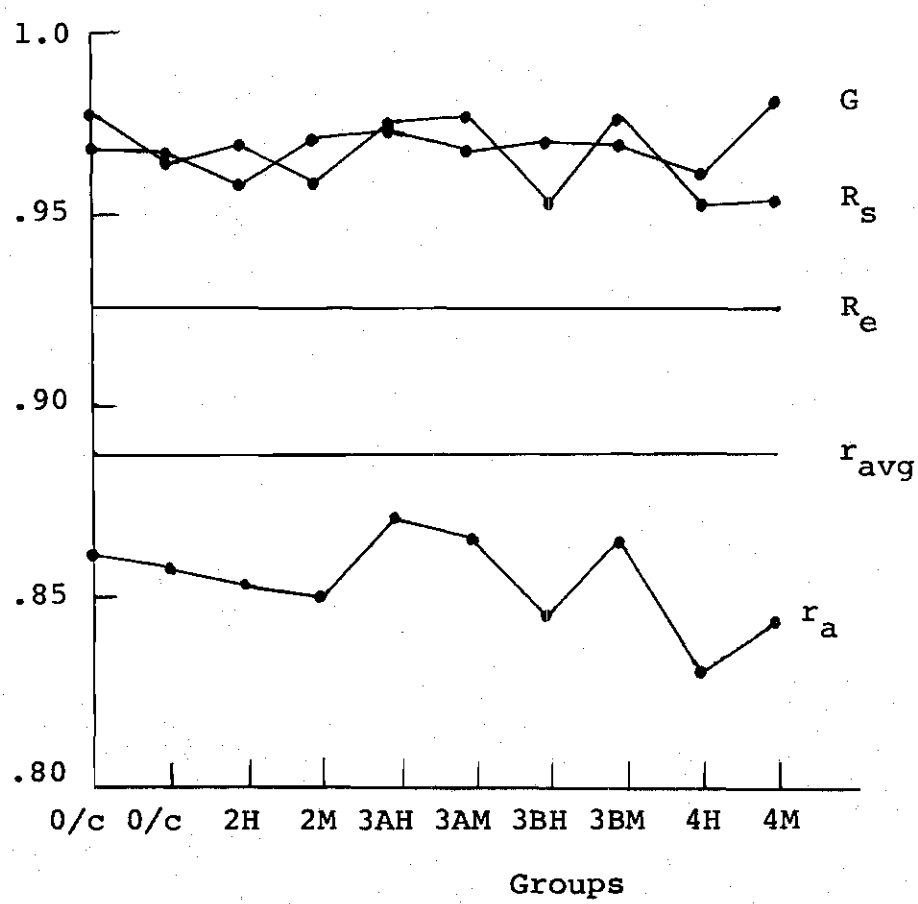

Figure 2. Mean r_a, R_s and G by Groups, Trials 41-80

receiving outcome feedback and those receiving ambiguous, categorical feedback. This hypothesis could not be rejected. These results and analysis are displayed in Table 3.

Table 3. Mean r_a, Outcome Feedback (Groups I and II) and Ambiguous, Categorical Feedback (Groups III-X)

Outcome Feedback	.860292
Amb., Cat. Feedback	.852755

Test of Hypothesis

H_O: $\mu_{r_{\sigma OC}} - \mu_{r_{\sigma Ac}} = 0$

H_1: $\mu_{r_{\sigma OC}} - \mu_{r_{\sigma Ac}} \neq 0$

t statistic = .584125009

α = P (rejecting a true Hypothesis) = .6

Fail to reject

The experiment had been divided into two sets of 40 trials. The first 40 trials without any feedback and the second 40 with the feedback appropriate to the group. An investigation was made of the difference in performance of the individual subjects between blocks. In all groups but two, 3 categories B with mild cutoffs and 4 categories with harsh cutoffs, no significant difference was found, the performance of the individuals deteriorated significantly at an α level of .1. These results are displayed in Table 4 and Figures 3 and 4.

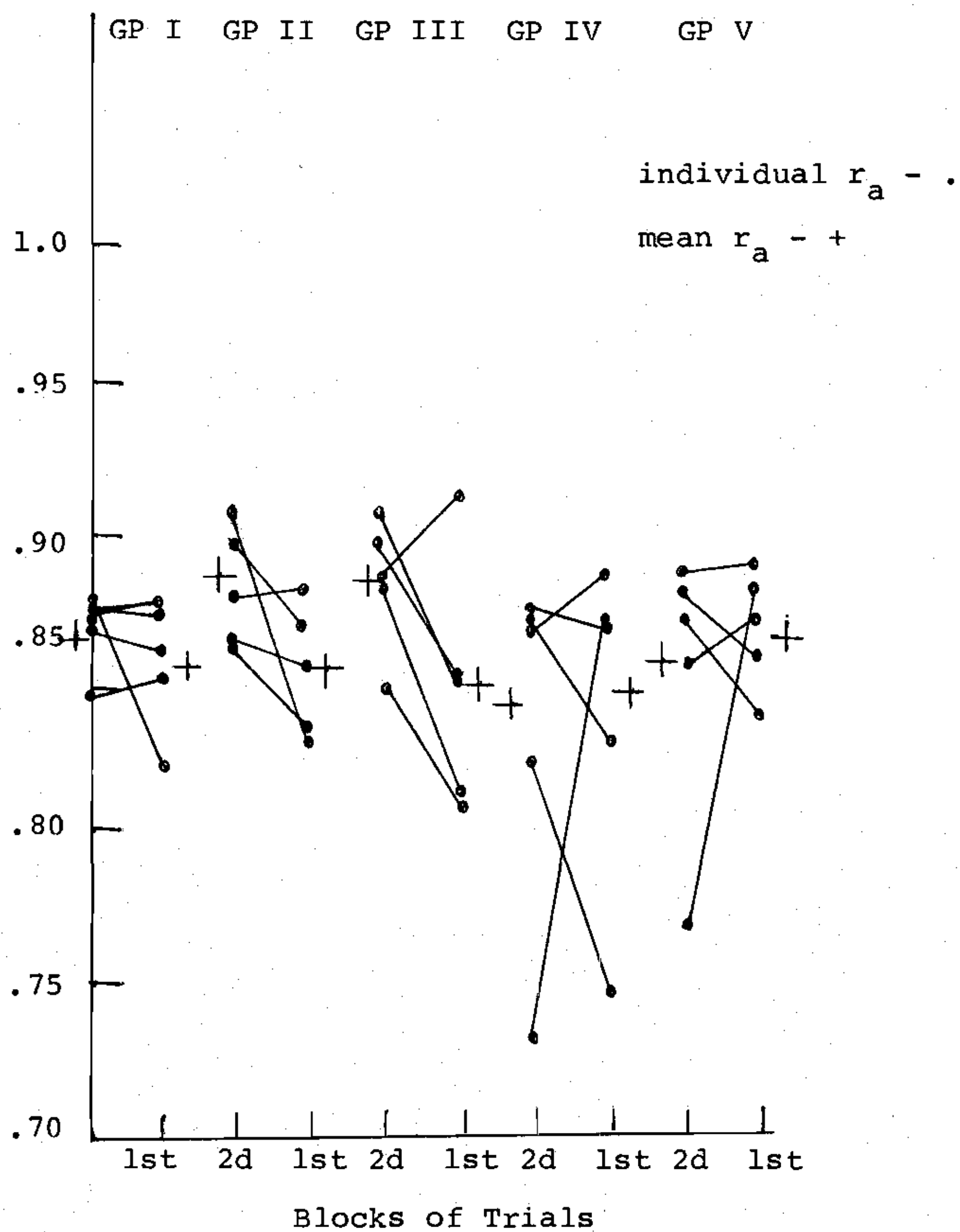

Figure 3. Individual Gain Scores Between First and Second Blocks by Groups

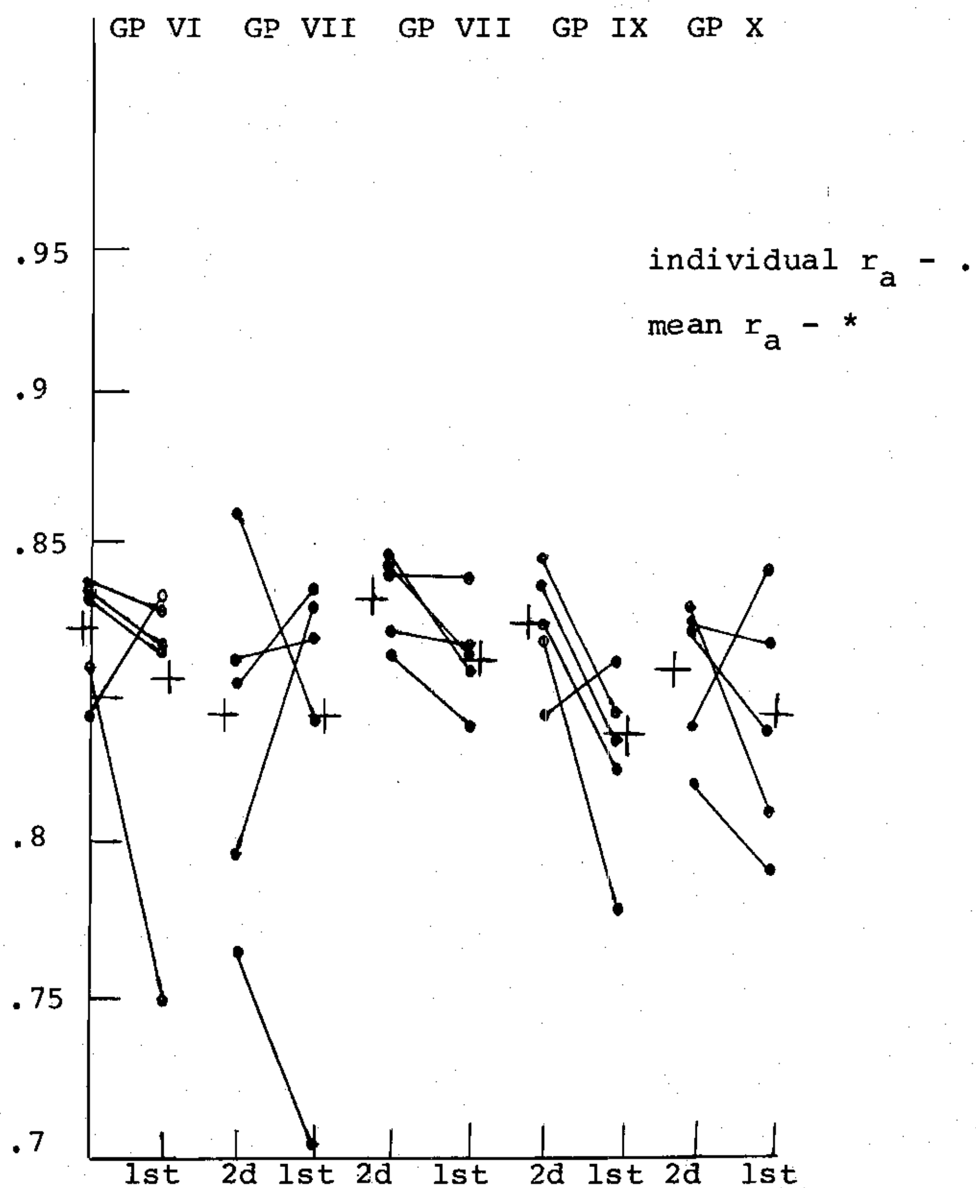

Figure 4. Individual Gain Scores Between First and Second Blocks, by Groups

Table 4. Gain Scores Between First and Second Blocks by Groups

Note: Gain Score = d = (r_a (trials 41-80))-(r_a (trials 1-40))

Group	Mean Gain Score	t_0^*	
I	-.042826	-1.082251104	N.S.
II	-.109059	-1.894530631	N.S.
III	-.131384	-2.063195986	N.S.
IV	.019108	.1727839121	N.S.
V	.044811	.5145227842	N.S.
VI	-.076976	-.9209508002	N.S.
VII	.00094	.0092294015	N.S.
VIII	-.080589	-2.4981900177	*
IX	-.158225	-2.653537005	*
X	-.052444	.709745504	N.S.

$t_0^* = /S_d/n{\sim}t_{\alpha/2}$, n-1 (Hines and Montgomery, 1972)

Additional Analysis

In addition to the previous analysis of the data a thorough investigation of the results took place.

ANOVA's were performed for the feedback trials (41-80) to study the effects of categories, cutoffs, and their interaction on R_s, G, and a non-correlational performance indicator, the summation of an individual's absolute deviations from the correct answer. In each case no treatment had any significant effect.

An ANOVA was performed to determine if there was any significant effects on performance across blocks of twenty trials. It was determined that the blocks of trials did have a highly significant effect on the performance of the subjects and that groups did not. Neither did block by group interaction. These results are displayed in Table 5.

Table 5. ANOVA on r_a by Individuals all Trials

Source of Variation	d.f.	M.S.	F ratio
Block	3	.7890689667	28.41761953****
Group	9	.0462949333	1.667270994
Block Group	27	.0134726963	.4852072172
Error	160	.0277668918	
Total	199		

This effect is confirmed in a comparison of the performance means across blocks. A difference in performance means is significant between the first 20 trials (no feedback) and the second 20 trials (no feedback) and again between the third 20 trials (feedback) and the last 20 trials (feedback). There was no significant difference in performance means between the second 20 trials (no feedback) and the third 20 trials (feedback). These results are displayed in Table 6 and Appendix H.

Table 6. Duncan Multiple Range Test Results

Block I	$\neq$	Block II
Block I	$\neq$	Block III
Block I	$\neq$	Block IV
Block II	$=$	Block III
Block II	$\neq$	Block IV
Block III	$\neq$	Block IV

Evaluations

The cutoff points utilized in the study were based in part on those of Rose (1974), for the groups in the study which performed under identical circumstances, and in part arbitrarily chosen to achieve either similarly harsh cutoffs or obviously milder cutoffs. The evaluations actually received by the subjects under these conditions are displayed in Appendix I.

Post Experimental Questionnaire Results

At the conclusion of the experiment each subject was verbally asked the four questions of the post experimental questionnaire (See Appendix F.

The methodology used by the subjects was in the vast majority (94%) some form of a weighting scheme. Only a small percentage (22%) said that they followed a simple scheme, eg. straight averaging, while the others utilized more exotic methods, eg. average of the two closest, weighted towards the third. These latter schemes evidenced some disregarding of one and even two of the information sources. The outcome feedback groups and the 4 category feedback groups evidenced more diversity in their methodology than the other groups. These results are displayed in Appendix J.

No difference is evidenced in the ability of the subjects across groups to identify the best and worst experts; although the latter identification seems slightly easier than the former. These results are displayed in Appendix K.

The majority (94%) of subjects said that they utilized the feedback to either determine the best expert or to determine the task structure. No other significant results are evidenced.

As expected the ambiguous, categorical feedback groups had different goals in their performance than the outcome feedback groups. No trend in this regard is evidenced among the former groups except that there seems to be some lack of direction for the 3 categories and the 4 categories groups. These results are displayed in Appendix L.

Summary

The data fail to support any of the hypotheses of this research. Secondary analysis indicated that no type of feedback had any effect on the performance of the subjects. A worsening of performance in inter trials was found, but this seems to be due exclusively to the number of trials undertaken. In short the present study represents a clear failure to replicate the findings of Rose (1974). No effect on performance of any of the feedback manipulations was found. This result will be discussed in the next chapter.

CHAPTER V

CONCLUSIONS AND RECOMMENDATIONS

Purpose

The purpose of this chapter is to state the conclusions of this research and to recommend possible future topics of research in multiple-cue probability learning studies.

Conclusions

The major hypotheses of this research were not confirmed. The achievement levels attained by all groups in the experiment were not significantly different; nor was the performance in the feedback trials significantly different from that in the no feedback trials.

These findings imply that performance in this multiple-cue probability learning task is not significantly effected by the feedback given. They also imply that the use of ambiguous, categorical feedback, rather than traditional outcome feedback, does not cause a significant deterioration in performance. Additionally neither outcome feedback nor ambiguous, categorical feedback significantly changes the performance of an individual from that under a no feedback condition in a multiple-cue probability learning task.

The manipulation of ambiguous, categorical feedback by changing the severity of the evaluative cutpoints, used in determining the feedback, is ineffectual as a mechanism of improving performance in the task. The lack of a significant effect reported in this study was indicated under both a relatively harsh evaluative criterion and a relatively mild evaluative criterion.

These results were in direct conflict with the findings of Rose (1974) who concluded that ambiguous, categorical feedback had a depressant effect on performance as the number of categories of feedback increased and that ambiguous, categorical feedback was inferior to no feedback at all. Further, no support was found for Rose's hypothesis that a lessening of the severity of the evaluative cutpoints would raise achievement level.

More generally, the negative findings in the present study run counter to a considerable body of work in the multiple-cue probability learning literature which broadly indicate that feedback does, in general, effect performance. For example, the provision of outcome feedback has been widely found to enhance performance (Azuma and Cronbach, 1966; Dudycha and Naylor, 1966; Todd and Hammond, 1965); feedback on problem structure, response characteristics, or both has been shown to have major effects (Hammond and Boyle, 1970; Todd and Hammond, 1965); error-prone feedback (Connolly and Miklausich, 1974) and other types of feedback (Castellan, 1974) have been used in laboratory studies, and have been generally shown to impact performance. The present findings are thus in conflict both with common sense and with careful empirical data, in their general suggestion that feedback does not matter. The most plausible explanation for the present findings would thus appear to be an experimental artifact peculiar to this study. Some possible artifacts are discussed below.

Vroom (1964) hypothesized that a person is motivated to perform effectively in terms of the relative strength of the forces acting on him to exert different levels of effort. These forces depend on (1) the strength of his preference for effective performance or ineffective performance, (2) his expectations regarding the consequences of different levels of effort on the attainment of effective or ineffective performance. Although experimentor observation and the Post Experimental Questionnaire of the experiment did not detect any lack of motivation among the subjects an indifference among the subjects regarding their performance on the problem and a lack of expectations regarding the consequences of their efforts was possible and in the latter aspect probably true.

These possibilities would seem to call for a replication of the study with perhaps some sort of incentive for effective performance. Studies have shown that knowledge of results by itself does not motivate performance directly, only by effecting the nature of the goals which individuals set on the task (Locke and Bryan, 1969).

The verbal context of this experiment was, in addition to the feedback, to be completely specified in objective terms in order to provide a controlled environment for the subjects (Hoffman, 1960). Simon (1947) hypothesized that executives perceive those aspects of a situation which relate specifically to the activities and goals of their departments and when presented with a complex stimulus he perceives in this stimulus what he is ready to perceive. Einhorn (1971) suggests that familiarity with a decision task will play a part in the decision strategy. Goldberg (1968) relates a study which demonstrates that prior expectations of the relationships between the cues and the distal variable of a problem can lead to faulty observation and inference. Perhaps the complexity of the verbal context of this problem, the unfamiliarity of the subjects with the stock market, or the development of prior expectations due to the verbal context, in regard to the cues and the distal variable, by the subjects led to the utilization of a particular strategy which was feedback resistant.

This could possibly have resulted in the study's surprising results. However as evidenced by both the Post Experimental Questionnaire and an analysis of the subjects' response sheets considerable shifting of strategies were in evidence. Hence this explanation does not seem plausible.

Another possible explanation for the poor performance evidenced across groups is that because of the cue sampling characteristics, it was quite possible for a subject to have had an appropriate model and yet occasionally receive bad evaluations. These evaluations could lead to abandonment of a good model and cause the subject to experiment with one possibly not as good. This is evidenced by a comparison of the performance of the subjects with a hypothetical averaging model's performance. In no case was a group's mean performance level as good as the averaging model's level (See Figure 1).

However the performance of the subjects in the no feedback trials was not significantly better than the performance during the feedback trials suggesting that the subject's original model was not significantly better than the subsequent ones. Additionally this abandonment and search problem would not seem to have been in evidence to the same extent under the mild evaluative conditions as under the harsh cutoffs yet the problem seems to exist under both.

An examination of the R_e for the blocks of trials reveals a change of only .002 (.926 for trials 1-40 to .928 for trials 41-80). In fact the problem was structured to avoid just such a dysfunctional cue characteristic. Another possible cue characteristic which could have lead to the study's results would be the existence of extreme cue values at critical points during the trials.

Several types of possible extremities were identified by the experimentor. These are: Case 1a - one cue is of a different sign than the other two cues; Case 1b - one cue is of a different sign than the other two cues with a difference in magnitude between the extreme cues of greater than 10 points; Case 2a-f - a difference of magnitude between the smallest and largest cues of greater than 10, 20, 30, 40, 50, and 60 points.

An examination of the trials (Trials 41-80) revealed that 42.5% (17 of 40) trials were Case 1a; 37.5% (15 of 40) were Case 1b; 62.5% (25 of 40) were Case 2a; 35% (14 of 40) were Case 2b; 12.5% (5 of 40) were Case 2c; 7.5% (3 of 40) were Case 2d; and only Trial 50 (1.25% - 1 of 40) was Case 2e and Case 2f.

Cases 1a, 1b, 2a and 2b were ruled out as possibilities since such large percentages seem to portray the essential probabilism of the task rather than any extremism. The investigation concentrated on the remaining cases, ie. Cases 2c-2f.

In these remaining possibilities a reasonably good model, strict averaging, would result in at least a bad evaluation in 80%, 100%, 100% and 100% of the trials. This would seem to indicate the possibility of dysfunctional feedback if the subject did have a reasonable good model prior to a particular trial.

However an examination of actual subject performance revealed that these particular trials did not result in any distinguishable change in strategy. Hence it must be concluded that the subjects' performances are not a direct result of the extremeness of any particular trials but rather a true reflection of the entire task's effect.

The implication of these conclusions may be beneficial in the search for an information system which can be utilized in a real world decision environment. They suggest that relatively poor feedback can be as useful to a decision maker as precise feedback.

These conclusions seem to run counter, to some extent, to past investigations for not only were the judgment policies executed in an inconsistent manner, as evidenced by the subjective questionnaire, but that the policies themselves were incorrect. A simple averaging scheme outperformed every group's performance.

The conclusions highlight Slovic's (1966) hypothesis with regard to cue consistency. The sampling error of the experiment seems to effect not only the sujbect's feel for the intrinsic validity of the cues but also sporadically results in a disagreement of the cues in their implications for the attributes being judged.

Lastly the conclusions reached are counter to the basic thrust of past multiple-cue probability learning research in that the outcome feedback groups evidence no significant learning over the last 40 trials.

Recommendations

It could be as Yntema and Torgerson (1961) suggest that "good judgment may turn out to be simpler matter than we would like to think it is". But if feedback does not signifiantly affect the decision maker, what does. More research must be conducted to investigate the implications of this study.

The use of one-half of the trials of this experiment with no feedback served no purpose other than as a control against which to compare the feedback trials. Future studies along these lines should abandon this approach and utilize only feedback trials for a more precise examination of the effects of an information structure.

A replication of this study with more severe and or milder evaluative cutoffs also seems futile. Possibly the evaluation of a response should be made based on a criterion of not how close the answer is in absolute terms but rather in some scaled terms dependent on the magnitude of the cues. This possibly could result in a problem which would seem to not only have more consistent cues, but would enhance the subject's feel for the intrinsic validity of the cues and their reliability. Although extremeness of the cues did not seem to have any direct effect on the results of the study perhaps another approach to this same problem would be the elimination of those trials for which it is a near impossibility for a subject to be anywhere near the correct answer, independent of the accuracy of his model or the consistency by which it is followed.

A more fruitful approach also may be a change in the type of feedback utilized. Perhaps changing from descriptive, common adjectives to defined quantitative or qualitative terms would produce feedback effects.

Perhaps the particular group of subjects utilized in the study, although extremely homogeneous, had detrimental effects on the results. As pointed out in the conclusions section of this chapter the motivation of the group was in question. A replication of this study with a group of subjects properly motivated would be highly desirable.

In summary then this research fails to determine any conclusive effects of the information structure on an individual in a judgment task. The implications of the study add to the body of knowldege regarding what it is a man does, what are the characteristics of his performance but not to how he acquires these characteristics. They clarify the vital questions of how people learn and grow.

This study also adds to the small body of empirical facts regarding behavior of individuals in an organizational setting by showing the seeming lack of influence of the environment on decision making. Thus perhaps it is a start in the attempt to find solutions to such problems that exist in the utilization of the theories of Management By Objectives and in the field of Research and Development Project Selection and Resource Allocation.

APPENDICES

APPENDIX A

A MATHEMATICAL DEVELOPMENT OF THE LENS MODEL

Assuming all variables are standardized, let

$$\hat{Y}_e := \beta_{e1}x_1 + \dots + \beta_{en}x_n$$

$$\hat{Y}_s := \beta_{s1}x_1 + \dots + \beta_{sn}x_n$$

where the β's are multiple linear regression weights for predicting the variables Y_e and Y_s from cue variables $x_1,\dots,x_n$ and; $\hat{Y}_e$ is the best linear prediction of Y_e from the cues and; $\hat{Y}_s$ is the best linear prediction of Y_s for the cues.

If Z_e and Z_s are defined as the residual errors for each variable then

$$Y_e = \hat{Y}_e + Z_e$$

$$Y_s = \hat{Y}_s + Z_s$$

Now VAR $(\hat{Y}_e) = R_e^2$ and VAR $(\hat{Y}_s) = R_s^2$

Where R_e^2 and R_s^2 are the squares of the multiple correlations involved and

VAR $(Z_e) = 1 - R_e^2$ and VAR $(Z_s) = 1 - R_s^2$ then

r_a, the correlation between Y_e and Y_s, can be shown to be

r_a = COV $(Y_e Y_s)$ = COV $(\hat{Y}_e \hat{Y}_s)$ + COV $(Z_e Z_s)$ and if correlation $(\hat{Y}_e \hat{Y}_s) = G$ and correlation $(Z_e Z_s) = C$ then

$$r_a = GR_eR_s + C\sqrt{1 - R_e^2}\ \sqrt{1 - R_s^2} \qquad \text{(Tucker, 1964)}$$

R_e - the linear predictability of the distal variable; the correlation between the distal variable and the linear model of the distal variable.

R_s - the linear predictability of the decision maker; the correlation between the decision maker's predictions and the linear model of the decision maker.

C - the nonlinear component of judgmental accuracy; the correlation between the residual values of the distal variable and the residual values of the decision maker's predictions after linear components in both have been removed.

G - a measure of the subject's "knowledge" (Hammond and Summers, 1972) of the "matching index" (Slovic and Lichtenstein, 1971); the correlation between the predicted response from the linear model of the decision maker and those from the linear model of the distal variable. (Rose, 1973)

when $C \cong 0$ then $r_a = GR_eR_s$ ie. with a linear task system the achievement correlation becomes simply a product of knowledge G the subject has of the task; R_e the predictability of the environment and R_s the predictability of the subject's response system.

APPENDIX B

CUE, DISTAL VARIABLE GENERATION PROGRAM

```
          DIMENSION Y(500),X(11,500),XT(500,11)
          SIGER(11)
          DIMENSION CV(11,11),IPUN(2),XBAR(200)
          PCOR(570)
          K=0
          READ(5,5)NSET
  5       FORMAT(12)
          READ(5,8) (IPUN(I),I=1,NSET)
  8       FORMAT(20I1)
        9 FORMAT (213)
          READ(5,9)L,NUM
          IP=L+1
          READ(5,1)YMU,SIG
          READ(5,1) (SIGER(I),I=1,L)
        1 FORMAT ( )
          DO 400 IM=1,NSET
          DO 20 J=1,NUM
          Z=0
          DO 10 I=1,12
 10       Z=Z+RNI(K)
          Z=Z-6
 20       Y(J)=YMU+SIG*Z
          DO 40 I=1,L
          DO 40 J=1,NUM
          EPS=0
          DO 30 N=1,12
 30       EPS=EPS+RNI(K)
          EPS=EPS-6
          EPS=EPS*SIGER(I)
 40       X(I,J)=Y(J)+EPS
          IPAS=IPUN(IM)
          IF (IPAS.EQ.O) GO TO 120
          DO 100 J=1,NUM
100       WRITE(1,110) (X(I,J),I=1,L),Y(J)
110       FORMAT(9F8.0)
120       DO 125 J=1,NUM
125       WRITE(6,130) J,(X(I,J),I=1,L),Y(J)
130       FORMAT(1X,I3,2X,11(F4.0,2X))
          DO 150 J=1,NUM
150       X(IP,J)=Y(J)
          DO 160 I=1,IP
          DO 160 J=1,NUM
```

```
160   XT(J,I)=X(I,J)
      CALL CORAN(XT,NUM,IP,1,O,XBAR,CV,PCOR
      500,11)
      DO 200 I=1,IP
200   WRITE(6,210) (CV(I,J),J=1,IP)
210   FORMAT(1X,6E13.6)
400   CONTINUE
      STOP
      FUNCTION RNI(K)
      IF(K.NE.O) GO TO 11
      KA=100933
      KB=5**7
      K=1
      GO TO 15
 11   KC=KD
 15   KC=KA*KB
      KD=MOD(KC,2**17)
      D=KD
      XX=D/(2.0**17)
      RNI=XX
      RETURN
      END
```

APPENDIX C

DISPLAY SHEETS

Cues

Trial	Expert 1	Expert 2	Expert 3	Distal (not shown)
1.	7	10	3	2
2.	-16	-9	-6	-4
3.	3	-7	-5	-2
4.	7	-13	-16	-9
5.	26	-2	-13	0
6.	20	8	14	7
7.	3	12	11	7
8.	0	-5	-5	3
9.	-31	-22	-36	-21
10.	-15	-18	-3	-11
11.	-8	-7	-21	-14
12.	-15	5	-1	5
13.	34	15	4	19
14.	-28	-12	-18	-8
15.	-1	-5	0	-3
16.	-6	-8	15	-1
17.	-13	-16	-4	-9
18.	7	-4	-1	-4
19.	19	15	13	20

Trial	Expert 1	Expert 2	Expert 3	Distal (not shown)
20.	6	-4	-16	-6
21.	-3	-10	-27	-8
22.	10	-16	-1	-11
23.	40	10	15	13
24.	50	13	-2	9
25.	-29	-12	-13	-10
26.	1	2	-11	-1
27.	30	14	10	11
28.	43	16	20	20
29.	-22	-17	-16	-11
30.	15	-4	1	1
31.	4	-10	-11	-9
32.	-1	-9	24	-7
33.	8	10	14	9
34.	-12	-3	-4	4
35.	-44	-7	-11	-8
36.	13	8	0	6
37.	-3	-2	-4	1
38.	8	15	0	0
39.	-19	2	23	4
40.	-5	2	-1	1
41.	7	0	2	2
42.	26	3	-7	1
43.	12	7	14	8
44.	-36	-21	-7	-14

Trial	Expert 1	Expert 2	Expert 3	Distal (not shown)
45.	11	-2	-3	-1
46.	7	5	-7	2
47.	3	6	10	6
48.	6	-13	-26	-13
49.	-2	8	14	8
50.	-59	-11	6	-17
51.	41	12	-2	6
52.	-3	2	8	10
53.	-7	15	-7	9
54.	24	0	11	7
55.	43	35	39	27
56.	1	19	5	14
57.	-4	-2	5	0
58.	-11	13	19	1
59.	33	25	12	19
60.	-12	-2	3	-1
61.	21	-6	-8	1
62.	8	10	3	6
63.	-4	2	4	7
64.	10	11	12	11
65.	-20	-3	-14	-1
66.	4	2	-7	6
67.	-3	-9	-11	-14
68.	-2	-1	-6	2
69.	27	2	3	8

Trial	Expert 1	Expert 2	Expert 3	Distal (not shown)
70.	-17	-14	-15	-12
71.	-45	-3	-2	-4
72.	14	4	8	5
73.	-28	-6	-17	-11
74.	-5	2	-10	1
75.	-3	-11	-2	-3
76.	8	-4	-16	-2
77.	6	8	5	10
78.	-4	-15	-12	-14
79.	22	9	22	9
80.	-10	0	-7	-7

APPENDIX D

SUBJECT RESPONSE SHEETS

(All Groups)

Trial	Expert 1	Expert 2	Expert 3	Est.
1.				
2.				
3.				
.				
.				
.				
40.				

(Groups I and II)

41.	Expert 1	Expert 2	Expert 3	Correct Answer	Est.
42.					
.					
.					
.					
80.					

(Groups III and IV)

Trial	Expert 1	Expert 2	Expert 3	Feedback	Est.
41.				G B	
42.				G B	
.					
.					
.					
80.				G B	

G - Good
B - Bad

(Groups V and VI)

Trial	Expert 1	Expert 2	Expert 3	Feedback	Est.
41.				VG G B	
42.				VG G B	
.					
.					
.					
80.				VG G B	

VG - Very Good
G - Good
B - Bad

(Groups VII and VIII)

Trial	Expert 1	Expert 2	Expert 3	Feedback	Est.
41.				G B VB	
42.				G B VB	
.					
.					
.					
80.				G B VB	

G - Good
B - Bad
VB - Very Bad

(Groups IX and X)

Trial	Expert 1	Expert 2	Expert 3	Feedback	Est.
41.				VG G B VB	
42.				VG G B VB	
.					
.					
.					
80.				VG G B VB	

VG - Very Good
G - Good
B - Bad
VB - Very Bad

APPENDIX E

INSTRUCTIONS TO SUBJECTS

You are about to participate in an experiment to measure learning. Make your decisions based on the information provided to you. The information, in the form of numbers represents daily stock market predictions ie. the Dow Jones Industrial Average change by three local "experts" eg. stock broker. Assume that each numbered trial is a new day.

I will display each day's prediction as I read them out loud. As each day's predictions are read, please record them and your prediction for that day on your answer sheet. Make your predictions solely on the information before you. When 40 trials have been completed please put down your pencils and a break will be given.

<u>Break</u>

(the above instructions were given to all groups)

Instructions for subjects in Groups I and II

Please look at your response sheets for trials 41-80. You will notice that a column is titled "correct answer". After each trial unlike those conducted before the break, I will feed back to you the actual market change, mark this down in the space provided. You will have an additional piece of information with which to make your predictions.

Instructions to subjects in Groups III - X

Please look at your response sheets for trials 41-80. You will notice that a column titled "Feedback" is provided. In this column you will notice letters provided for my evaluation of your prediction on each trial. The legend for the letters is provided at the bottom of the column. During these trials I will evaluate your prediction after each trial. The evaluation is based on how close your prediction is to the actual market change. You now have an additional piece of information with which to make your prediction.

APPENDIX F

POST EXPERIMENTAL QUESTIONNAIRE

Now that you have completed all 80 trials please answer the following questions:

(1) How did you make your predictions?

(2) Did you feel any expert was better than the others? If so, which one?
Did you feel any expert was worse than the others? If so, which one?

(3) How did you utilize the feedback?

(4) What was your goal?

Note: These instructions were given at the conclusion of the test period to all subjects

APPENDIX G

DATA ANALYSIS PROGRAM

```
      DIMENSION X(200 ,30),W(200),SBAR(30)
      A(30,30),SIG(30),IVAR(29)
    1 B(29),SB(29),R(200),RESP(200,4)
      IT=1
      IPC=0
      ICOUNT=0
      MAXN=200
      MAXNP1=30
  100 IF (ICOUNT) 11,11,200
  200 READ (5,205) (X(I,NP1),I=1,N)
  205 FORMAT ( )
      GO TO 198
   11 READ (5,1000,END 79) ITEST,N,NP1,IW
      LIMIT,NS,EFIN,EFOUT
 1000 FORMAT (6I3,F4.0,F4.0)
      NSP=0
      ICOUNT=0
      WRITE (6.1011)
      WRITE (6,1012) ITEST,N,NP1,IW,LIMIT
      EFIN,EFOUT
 1011 FORMAT (11H1 TEST CASE,7X,1HN,5X,3HNP1,6X
      2HIW,4X,5HLIMIT,4X
    1 4HEFIN,3X,5HEFOUT)
 1012 FORMAT (I8,I11,3I8,F9.2,F8.2//)
      NP=NP1-1
      WRITE (6,1013)
 1013 FORMAT (10H RAW DATA )
      DO 1 I=1,N
      READ (5,1001) (X(I,J),J=1,NP1)
    1 WRITE (6,1014) (X(I,J),J=1,NP1)
 1014 FORMAT (9F13.3)
 1001 FORMAT (9F8.0)
  198 IF (W) 2,2,3
    3 READ (5,1001) (W(I),I=1,N)
      WRITE (6,1016) (W(I),I=1,N)
 1016 FORMAT (//19HO WEIGHTING FACTORS/(9F13.3))
    2 IND=0
      ISTEP=-1
    7 CALL RESTEM(X,N,NP1,MAXN,MAXNP1,W,IW,EFIN
      EFOUT,EBAR,A,SIG,CONST
    1 NVAR,FLEVEL,SY,NOIN,IVAR,B,SB,R,IND)
      IF (IND) 13,12,13
```

```
   13 ISTEP=ISTEP+1
      IF (ISTEP) 4,5,4
    5 WRITE (6,1002)
 1002 FORMAT(1H1,//26HO CORRELATION COEFFICIENTS)
      MM=NP-1
      DO 6 I=1,MM
      II=I+1
    6 WRITE (6,1003) (I,J,A(I,J),J=II,NP)
 1003 FORMAT (3(3H   I2,6H  VS  I2,5H  =
      F10.6,1X))
      WRITE  (6,1004) SY
 1004 FORMAT  (22HO STANDARD ERROR OF Y
      F12.6///)
      GO TO 7
    4 WRITE (6,1005) ISTEP
 1005 FORMAT (11HO STEP NO. I3)
      IF (NVAR) 8,8,9
    8 NVAR=-NVAR
      WRITE (9,1006) NVAR
 1006 FORMAT (5X,19H VARIABLE ENTERING .I3)
      GO TO 10
    9 WRITE (6,1007) NVAR
 1007 FORMAT (5X,19H VARIABLE ENTERING I3)
   10 WRITE (6,1008) FLEVEL,SY,CONST,(IVAR(I)
      B(I),SB(I),I 1,NOIN)
 1008 FORMAT (5X,7H FLEVEL,F13.6/5X,20H STANDARD
      ERROR OF Y,F13.6/5X,9H
     1CONSTANT,F13.6//15X,46H VARIABLE
      COEFFICIENT STD ERROR OF CO
     1EFF,/(17X,3H X I2,F16.5,F18.5))
      IF (ISTEP-LIMIT) 7,14,14
   14 IND=-1
      GO TO 7
   12 WRITE (6,1009)
 1009 FORMAT (30H1 PREDICTED VS. ACTUAL RESULTS
      /8H OB. NO.,8X,7H ACTUAL
      110X,10H PREDICTED,9X,10H DEVIATION)
      DO 20 I=1,N
      DEV=X(I,NP1)-R(I)
   20 WRITE (6,1010) I,X(I,NP1),R(I),DEV
 1010 FORMAT (I5,3(F18.4))
      ICOUNT=ICOUNT+1
      IF (ICOUNT-1) 45,45,55
   45 DO 50 I=1,N
      RESP(I,1)=X(I,NP1)
   50 RESP(i,2)=R(I)
      GO TO 200
   55 DO 60 I=1,N
      RESP(I,3)=X(I,NP1)
   60 RESP(I,4)=R(I)
      NSP=NSP+1
```

```
      WRITE (6,65) NSP
   65 FORMAT(1H1,19HDATA FOR SUBJ. NO. ,I3)
      CALL CORREL(RESP,N,4,XBAR,200,4)
      IF (NSP.GE.NS) GO TO 11
      GO TO 100
   79 STOP
      END
```

APPENDIX H

DUNCAN MULTIPLE RANGE TEST ON MEAN r_a, BY BLOCKS

$s_e = .0221305106$
$n_2 = 160$

	2	3	4
Significant Range	3.69	4.18	4.48
Least Significant Range	.0816615841	.0925055343	.0991446875

$Z_I = 1.42749184$	I vs. IV	.2998019 >	.0991446875
$Z_{II} = 1.29354686$	I vs. II	.13394498 >	.0925055343
$Z_{III} = 1.33705296$	I vs. III	.09043888 >	.0816615841
$Z_{IV} = 1.12768994$	III vs. IV	.20936302 >	.0925055343
	III vs. II	.0435061 <	.0816615841
	II vs. IV	.16585692 >	.0816615841

APPENDIX I

INDIVIDUAL AND GROUP EVALUATIONS

Group III (2 cat. harsh)	G		B	
subject 1	17	(42.5%)	23	(57.5%)
subject 2	12	(30%)	28	(70%)
subject 3	10	(25%)	30	(75%)
subject 4	13	(32.5%)	27	(67.5%)
subject 5	15	(47.5%)	25	(62.5%)
$\bar{X}$	13.4	(33.5%)	26.6	(66.5%)
Group IV (2 cat. mild)				
subject 1	22	(55%)	18	(45%)
subject 2	21	(52.5%)	19	(47.5%)
subject 3	19	(47.5%)	21	(52.5%)
subject 4	24	(60%)	16	(40%)
subject 5	23	(57.5%)	17	(42.5%)
$\bar{X}$	21.8	(54.5%)	18.2	(45.5%)

Group V (3 cat. A harsh)	VG		G		B	
subject 1	11	(27.5%)	2	(5%)	27	(67.5%)
subject 2	10	(25%)	3	(7.5%)	27	(67.5%)
subject 3	8	(20%)	7	(17.5%)	25	(62.5%)
subject 4	10	(25%)	1	(2.5%)	29	(72.5%)
subject 5	10	(25%)	1	(2.5%)	29	(72.5%)
$\bar{X}$	9.8	(24.5%)	2.8	(7%)	27.4	(68.5%)
Group VI (3 cat. A mild)						
subject 1	17	(42.5%)	5	(12.5%)	18	(45%)
subject 2	14	(35%)	9	(22.5%)	17	(42.5%)
subject 3	19	(47.5%)	5	(12.5%)	16	(40%)
subject 4	14	(35%)	3	(7.5%)	23	(57.5%)
subject 5	17	(42.5%)	3	(7.5%)	20	(50%)
$\bar{X}$	16.2	(40.5%)	5	(12.5%)	18.8	(47%)

	G		B		VB	
Group VII (3 cat. B harsh)						
subject 1	11	(27.5%)	12	(30%)	17	(42.5%)
subject 2	12	(30%)	8	(20%)	20	(50%)
subject 3	12	(30%)	15	(37.5%)	13	(32.5%)
subject 4	9	(22.5%)	12	(30%)	19	(47.5%)
subject 5	9	(22.5%)	6	(15%)	25	(62.5%)
$\bar{X}$	10.6	(26.5%)	10.6	(26.5%)	18.8	(47%)
Group VIII (3 cat. B mild)						
subject 1	23	(57.5%)	12	(30%)	5	(12.5%)
subject 2	20	(50%)	15	(37.5%)	5	(12.5%)
subject 3	20	(50%)	10	(25%)	10	(25%)
subject 4	24	(60%)	11	(27.5%)	5	(12.5%)
subject 5	18	(45%)	12	(30%)	10	(25%)
$\bar{X}$	21	(52.5%)	12	(30%)	7	(17.5%)

Group IX (4 cat. harsh)	VG		G		B		VB	
subject 1	8	(20%)	5	(12.5%)	11	(27.5%)	16	(40%)
subject 2	8	(20%)	5	(12.5%)	7	(17.5%)	20	(50%)
subject 3	9	(22.5%)	5	(12.5%)	8	(20%)	18	(45%)
subject 4	7	(17.5%)	2	(5%)	12	(30%)	19	(47.5%)
subject 5	6	(15%)	1	(2.5%)	11	(27.5%)	22	(55%)
$\bar{X}$	7.6	(19%)	3.6	(9%)	9.8	(24.5%)	19	(47.5%)
Group X (4 cat. mild)								
subject 1	13	(32.5%)	8	(20%)	12	(30%)	7	(17.5%)
subject 2	14	(35%)	6	(15%)	12	(30%)	8	(20%)
subject 3	16	(40%)	11	(27.5%)	11	(27.5%)	2	(5%)
subject 4	16	(40%)	9	(22.5%)	6	(15%)	9	(22.5%)
subject 5	16	(40%)	10	(25%)	7	(17.5%)	7	(17.5%)
$\bar{X}$	15	(37.5%)	8.8	(22%)	9.6	(24%)	6.6	(16.5%)

by severity level:

Harsh	11.95 (29.875%)	28.05 (70.125%)
Mild	21.95 (54.875%)	18.05 (45.125%)

APPENDIX J

SUBJECTS' METHODS

Method	Description	Percentage of subjects utilizing method
A	averaging	.12
B	weighted averaging	.04
C	averaging, searching for a pattern	.04
D	averaging and median	.04
E	averaging if all cues close if not averaging of the two closet biased towards the extreme	.12
F	averaging of the two closet biased towards the extreme	.30
G	averaging of the two closet	.10
H	following best previous cue or averaging biased towards the extreme	.02
I	weighted average biased towards previous best	.06
J	estimated mean of a hypothetical distribution	.02
K	following second expert	.06
L	various weighting schemes	.06

Methods by Groups:

Group I	.2 - A	Group II	.2 - A	Group III	.2 - A
	.2 - C		.4 - E		.6 - F
	.2 - F		.2 - G		.2 - G
	.2 - I		.2 - K		
	.2 - K				
Group IV	.2 - A	Group V	.2 - D	Group VI	.4 - B
	.2 - D		.4 - F		.2 - F
	.6 - F		.2 - J		.2 - G
			.2 - L		.2 - H
Group VII	.2 - E	Group VIII	.4 - E	Group IX	.4 - A
	.4 - F		.4 - F		.2 - C
	.4 - L		.2 - I		.2 - E
					.2 - G
Group X	.4 - F				
	.2 - G				
	.2 - I				
	.2 - K				

Methods by Severity Level:

Harsh	.15 - A	.35 - F	Mild	.05 - A	.4 - F
	.05 - C	.1 - G		.1 - B	.1 - G
	.05 - D	.05 - J		.05 - D	.05 - H
	.1 - E	.15 - L		.1 - E	.1 - I
					.05 - K

Methods by Categories:

O/C		2 cat.		3 cat. A	
	.2 - A		.2 - A		.2 - B
	.1 - C		.1 - D		.1 - D
	.2 - E		.6 - F		.3 - F
	.1 - F		.1 - G		.1 - G
	.1 - G				.1 - H
	.1 - I				.1 - J
	.2 - K				.1 - L

3 cat. B		4 cat.	
	.3 - E		.2 - A
	.4 - F		.1 - C
	.1 - I		.1 - E
	.2 - L		.2 - F
			.2 - G
			.1 - I
			.1 - K

APPENDIX K

EXPERT IDENTIFICATION BY GROUPS

Group	Correctly picked the best expert	Correctly picked the worst expert
I	4 of 5	5 of 5
II	1 of 5	3 of 5
III	3 of 5	1 of 5
IV	2 of 5	1 of 5
V	2 of 5	3 of 5
VI	0 of 5	3 of 5
VII	1 of 5	2 of 5
VIII	2 of 5	2 of 5
IX	3 of 5	4 of 5
X	3 of 5	3 of 5

APPENDIX L

SUBJECT GOALS

	Goal Description	Percentage of subjects for which this was the goal
A	To be as close as possible or the correct answer	.30
B	Maximize feedback	.42
C	Determine pattern, method or best expert	.20
D	others or meaningless	.08

Goals by Groups:

Group I	A - 1.0	Group VI	A - .4	Group IX	B - .6
Group II	A - 1.0		B - .4		C - .2
Group III	A - .2		C - .2		D - .2
	B - .8	Group VII	B - .6	Group X	A - .2
Group IV	B - .4		C - .2		B - .4
	C - .6		D - .2		D - .4
Group V	B - .4	Group VIII	A - .4		
	C - .6		B - .6		

Goals by Severity Level:

Harsh	A - .05	Mild	A - .25
	B - .6		B - .45
	C - .25		C - .2
	D - .1		D - .1

Goals by Categories:

O/C	A - 1.0	3 cat. A	A - .2	3 cat. B	A - .2	4 cat.	A - .1
2 cat.	A - .1		B - .4		B - .6		B - .5
	B - .6		C - .4		C - .1		C - .1
	C - .3				D - .1		D - .3

BIBLIOGRAPHY

1. Baker, Norman R., C. Richard Shumway, P. Michael Maher, William E. Souder, and Robert D. Large, "Army Material Command Seminar on Value Measurement", instruction booklet, Northwestern University, 1970.

2. Baker, Norman R. and James R. Freeland, "Recent Advances in Research and Development Value Measurement and Project Selection Methods", meeting paper, 1972.

3. Beach, Lee Ray, "Multiple Regression as a Model for Human Information Utilization", Organizational Behavior and Human Performance, Volume 2, 1967, pp. 276-289.

4. Bjorkman, Mats, "Feedforward and Feedback as Determiners of Knowledge and Policy: Notes on a Neglected Issue", Scandanavian Journal of Psychology, Volume 13, 1972, pp. 152-158.

5. Brady, Rodney H., "MBO Goes to Work in the Public Sector", Harvard Business Review, March-April 1973, Volume 51, number 2, pp. 65-74.

6. Brandenberg, Richard G., "Project Selection in Industrial Research and Development: Problems and Decision Processes", Research Program Effectiveness, M.C. Yovits et al eds., 1966, pp. 13-40.

7. Castellan, N. John Jr., "The Analysis of Multiple Criteria in Multiple Cue Judgment Tasks", Organizational Behavior and Human Performance, 8, 1972, pp. 242-261.

8. Castellan, N. John Jr., "Comments on the Lens Model Equation and Analysis of Multiple Cue Judgment Tasks", Psychometrika, Volume 38, number 1, March 1973.

9. Castellan, N. John Jr., "The Effects of Different Types of Feedback in Multiple Cue Probability Learning", Organizational Behavior and Human Performance, 11, 1974, pp. 44-64.

10. Cetron, Marvin J., Joseph Martino, and Lewis Roepche, "The Selection of Research and Development Program Content - Survey of Quantitative Methods", IEEE Transactions on Engineering Management, March 1967.

11. Connolly, Terry, "The Diffuse Decision: An Integrating Focus", a meeting paper, 1972.

12. Connolly, Terry and Val Miklausich, "The Effects of Error Prone Outcome Feedback on Learning and Performance in a Multiple Cue Judgment Task", a meeting paper, 1974.

13. Cook, Doris M., "The Impact on Managers of Frequency of Feedback", Academy of Management Journal, September 1968, pp. 263-277.

14. Dean, Burton V. and Lewis A. Roepche, "Cost Effectiveness in Research and Development Organizational Resource Allocation", IEEE Transactions on Engineering Management, November 1969.

15. Dudycha, Arthur L., Myron G. Dimoff, and Linda W. Dudycha, "Choice Behavior in Dynamic Environments", Organizational Behavior and Human Performance, 9, 1973, pp. 328-338.

16. Dudycha, Linda W., and James C. Naylor, "Characteristics of the Human Inference Process in Complex Choice Behavior Situations", Organizational Behavior and Human Performance, 1, 1966, pp. 110-128.

17. Duncan, Daniel M., "MBO at the Operating Level", Supervisory Management, June 1973, pp. 11-14.

18. Einhorn, Hillel J., "Expert Measurement and Mechanical Combination", Organizational Behavior and Human Performance, 7, pp. 86-106, 1972.

19. Freeland, James Ross, "Conceptual Models of the Resource Allocation Decision Process in Hierarchical Decentralized Organizations", doctoral dissertation, Georgia Institute of Technology, 1973.

20. Goldberg, Lewis, "Simple Models or Simple Processes", American Psychologist, Volume 23, 1968, pp. 483-496.

21. Hammond, Kenneth R., "Computer Graphics as an Aid to Learning", Science, 1971, Volume 172, pp. 903-908.

22. Hammond Kenneth R., and David A. Summers, "Cognitive Control", Psychological Review, Volume 79, number 1, 1972, pp. 58-67.

23. Hestland, Richard R., "MBO: Lets Start Managing Objectives", Supervisory Management, 1973, Volume 18, number 1, pp. 2-10.

24. Hines, William W., and Douglas C. Montgomery, Probability and Statistics in Engineering and Management Science, New York: The Ronald Press Company, 1972.

25. Hoffman, Paul J., "The Paramorphic Representation of Clinical Judgment", Psychological Bulletin, Volume 57, number 2, 1960.

26. Howell, Robert A., "Managing by Objectives - A Three Stage System", Business Horizons, February 1970, pp. 41-45.

27. Hursch, Carolyn J., Kenneth R. Hammond, and Jack L. Hursch, "Some Methodological Considerations in Multiple Cue Probability Studies", Psychological Review, 1964, Volume 71, number 1, pp. 42-60.

28. Locke, Edwin A., and Judith F. Bryan, "The Directing Function in Task Performance", Organizational Behavior and Human Performance, 4, pp. 35-42, 1969.

29. McCMiller, Patrick, "Do Labels Mislead? A Multiple Cue Study, Within the Framework of Brunswick's Probabilistic Functionalism", Organizational Behavior and Human Performance, 6, 1971, pp. 480-500.

30. McNemar, Quinn, Psychological Statistics, New York: John Wiley and Sons, Inc., 1949.

31. Miklausich, Val., "The Impact of Error Prone Feedback on Learning in a Multiple Cue Decision Problem", master's thesis, Georgia Institute of Technology, 1973.

32. Moore, John R. Jr., and Norman R. Baker, "Computational Analysis of Scoring Models for Research and Development Project Selection", Management Science, Volume 16, number 4, December 1969.

33. Odiorne, George S., Management By Objectives, New York: Pitman Publishing Corporation, 1965.

34. Parrish, Leo Gillis Jr., "Short Term Effects on Employee Attitudes of the Introduction of Management By Objectives - An Empirical Study in a Local Government", doctoral dissertation, Georgia Institute of Technology, 1973.

35. Peterson, Cameron R., Kenneth R. Hammond, and David A. Summers, "Optimal Responding in MCPL", Journal of Experimental Psychology, 1965, Volume 70, number 3, pp. 270-276.

36. Rose, Donald Edward Jr., "The Impact of Evaluative Feed-

back on Learning and Performance in a Multiple Cue Decision Problem", master's thesis, Georgia Institute of Technology, 1974.

37. Schein, Edgar H., "Management Development as a Process of Inference", Industrial Management Review, May 1961, pp. 59-77.

38. Schench, E. Allen and James C. Naylor, "A Cautionary Note Concerning the Use of Regression Analysis for Capturing the Strategies of People", Educational and Psychological Measurement, 1968, 28, pp. 3-7.

39. Simon, H.A., Administrative Behavior, New York: MacMillan, 1947.

40. Slovic, Paul, "Cue Consistency and Cue Uitlization in Judgment", American Journal of Psychology, 1966, Volume 75, pp. 427-434.

41. Slovic, Paul and Sarah Lichtenstein, "Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment", Organizational Behavior and Human Performance, 6, 1971, pp. 649-744.

42. Todd, Frederick, J., and Kenneth R. Hammond, "Differential Feedback in Two Cue Multiple Cue Probability Learning Tasks", Behavorial Science, 1965, 10, pp. 429-435.

43. Tucker, Ledyard R., "A Suggested Alternative Formulation in the Developments by Hursch, Hammond and Hursch and Hursch, Hammond, and Todd, Psychological Review, 1964, Volume 71, number 6, pp.528-530.

44. Vroom, Victor H., Work and Motivation, New York: John Wiley and Sons, Inc., 1964, pp. 186-187.

45. Yntema, D. B., and Torgeson, W. S., "Man-Computer Cooperation in Decisions Requiring Common Sense", IEEE Transactions on Human Factors in Electronics, HFE-2, pp. 20-26, 1961.